Cassandra Crossing/ Archivismi: Cassandra attraverso i secoli

(568)—Cassandra non si accontenta, vuole arrivare più lontano e vuole sopravvivere non per decenni ma per secoli o millenni. Ce la può…

Cassandra Crossing/ Archivismi: Cassandra attraverso i secoli

Figure 1:

(568)—Cassandra non si accontenta, vuole arrivare più lontano e vuole sopravvivere non per decenni ma per secoli o millenni. Ce la può fare?

10 gennaio 2024—Nelle precedenti 10 puntate di *Archivismi* abbiamo descritto la prima *campagna di archiviazione*; quella della rubrica Cassandra Crossing su I*nternet Archive*. E' stato un lungo percorso, poiché siamo partiti dallo studio della struttura di Internet Archive, seguito la preparazione dei dati, realizzato qualche decina di righe di script per automatizzare il tutto, eseguito gli upload veri e propri, ed infine la ripulitura dei dati e la correzione degli errori nei metadati.

Oggi invece introdurremo la terza campagna di archiviazione di Cassandra Crossing.

"*Ohibò*—dirà qualcuno dei più informati 24 lettori—*la terza campagna? Ma dove ci hai raccontato la seconda?*"

Giustissimo, la seconda non l'ho raccontata perché è stata troppo facile e veloce.

La seconda campagna consisteva nell'archiviazione dei 106 video di ***Quattro Chiacchiere con Cassandra*** su *Internet Archive*. Cassandra ha deciso di non parlarne perché è appena

finita, ed ha richiesto solo 20 minuti di preparazione del foglio elettronico di bulk upload e circa un'ora di caricamento. E' pur vero che avevamo maturato una preziosa esperienza precedente, che i metadati inseriti sono elementari e che i dati di partenza erano già ben strutturati, ma una cosa così semplice e veloce non poteva meritare una pur breve esternazione di Cassandra. Per cui la butto li, andatevi a vedere il risultato, e passiamo davvero alla terza campagna di archiviazione, che ve lo anticipo, sarà ben più stuzzicante.

Dobbiamo però, come Cassandra vi ha ormai abituato, raccontare un po' di storia. Veramente assai più di un po', visto che non si tratta di partire dall'alba di Internet, nemmeno dall'alba dei computer, ma addirittura dall'alba della scrittura, il che vuol dire riavvolgere il nastro, così all'ingrosso, di 5 millenni abbondanti. E' da quella remota epoca che è giunto fino a noi il primo archivio di informazioni omogenee, scritto in caratteri cuneiformi su circa 4.000 tavolette di argilla. Se consideriamo la tavoletta di argilla come supporto informativo, potremmo dire che le tavolette di Uruk si sono rivelate molto durevoli, facendo impallidire tutti i moderni supporti informatici.

E' pur vero che innumerevoli altre tavolette di argilla non hanno superato, come le loro più famose 4000 colleghe, il lungo viaggio fino a noi, ma comunque l'efficacia del supporto rimane notevole.

I rotoli di pergamena si sono rivelati poco meno durevoli; i più antichi superano infatti di poco i duemila anni, e la durata "media" della pergamena, conservata in condizioni ideali, è stimata intorno ai mille anni.

Alcuni papiri sono giunti a noi dall'antico Egitto e quindi sono durati anche loro per millenni, ma in condizione estremamente particolari (tombe sigillate nel deserto). Nei climi europei ed in condizioni di conservazione ideali hanno invece una durata stimata intorno ai 300 anni. Vale la pena di notare che la scomparsa della pergamena come supporto per le informazioni è dovuto proprio all'avvento del papiro, più economico, più facile da scrivere, più leggibile ma meno durevole.

L'avvento della carta ha ulteriormente peggiorato le cose; se alcuni volumi del passato hanno superato molti secoli, tutta la produzione moderna di carta ha una durata limitata a pochi decenni, con casi estremi come certi tascabili degli anni '90 o la carta di giornale, che bastava lasciare al sole per vederla letteralmente sbriciolarsi. Colpa di addittivi chimici e sbiancanti, usati per migliorarne l'aspetto, e di processi di lavaggio inefficienti.

Possiamo riassumere che c'è stato un progresso continuo tra un supporto e l'altro che ha prodotto costi minori, prestazioni migliori e durate peggiori. D'altra parte sostituire supporti inorganici ed incombustibili con supporti organici e combustibili non poteva che peggiorare la durata delle informazioni ivi registrate.

In campo informatico non c'è una esperienza storica così lunga. Inizia solo dagli anni '50 del secolo scorso, con le schede perforate (e per inciso ne ho un pacchetto in perfetto stato di conservazione in un cassetto, perforate per la tesi nel 1980).

I supporti informatici, magnetici od ottici, hanno avuto performance assai meno brillanti. A parte l'obsolescenza tecnologica intrinseca delle periferiche di lettura/scrittura, divenute introvabili o non funzionanti, che rende illeggibili anche supporti che sarebbero ben conservati, persino i nastri magnetici ed i cd-rom, che vantavano durate di 30 anni, si sono in realtà rivelati molto più cagionevoli del previsto. Una campagna di trasferimento dati eseguita di persona da CD-R di meno di venti anni conservati in condizioni ideali ha portato a quasi il 10% di supporti con problemi più o meno gravi di lettura.

La triste verità è che lo sviluppo dell'informatica moderna ha sempre privilegiato la riduzione

del costo unitario dei supporti, la densità delle informazioni ivi registrate, la velocità di accesso alle informazioni stesse, senza porre una equivalente cura alla durata dei supporti stessi.

E questo può essere sufficiente per spiegare come mai la durata dei supporti, a partire dai 20–30 anni degli anni '60, non sia migliorata ma anzi sia semmai peggiorata. Non stiamo infatti parlando di sistemi dotati di ridondanza ed algoritmi di correzione; questi sistemi devono essere dinamici, consumano energia e sono soggetti comunque a problemi di sicurezza informatica e di scarsa resilienza alle catastrofi.

Quello che serve sono supporti che conservino in maniera affidabile le informazioni per la loro stabilità e durata intrinseche, ed in maniera completamente passiva, senza consumare energia, né direttamente, come una stringa di dischi in RAID che deve essere alimentata e funzionante per essere stabile, né indirettamente, a causa di processi produttivi costosi e/o la necessità di impianti attivi di conservazione, come condizionamento/riscaldamento per la stabilizzazione della temperatura.

E servono anche supporti in cui la rappresentazione dei dati non sia così "lontana" dalla percezione degli utenti. La maggior parte delle unità di lettura/scrittura di dati digitali producono supporti sui quali i dati sono impercettibili con mezzi normali, e possono essere rivelati solo con un particolare tipo di unità hardware.

Ambedue queste caratteristiche sono presenti nella soluzione che, attualmente, garantisce i tempi di conservazione più lunghi tra i prodotti disponibili sul mercato. E, curiosamente, ma forse non per caso, si tratta di una tecnologia abbastanza vecchia, a cui sono state apportati alcuni miglioramenti. Parliamo delle pellicole fotografiche "normali", cioè all'alogenuro di argento, ed in particolare di quella utilizzata dalla Piql, una azienda norvegese, insieme al macchinario per registrarvi informazioni digitali.

Il formato della pellicola, che è un prodotto commerciale, è il normale 35 millimetri, il supporto usato è un tipo di poliestere, e la gelatina e l'emulsione hanno ovviamente caratteristiche particolari. La durata di questa pellicola, opportunamente impressionata e sviluppata, è stimato poter arrivare a 500 anni, conservata a temperatura ambiente ed in condizioni ottimali.

La scrittura dei dati sulla pellicola, che alla fine è comunque una normale pellicola fotografica, può avvenire in vari modi, sia visuali che codificati.
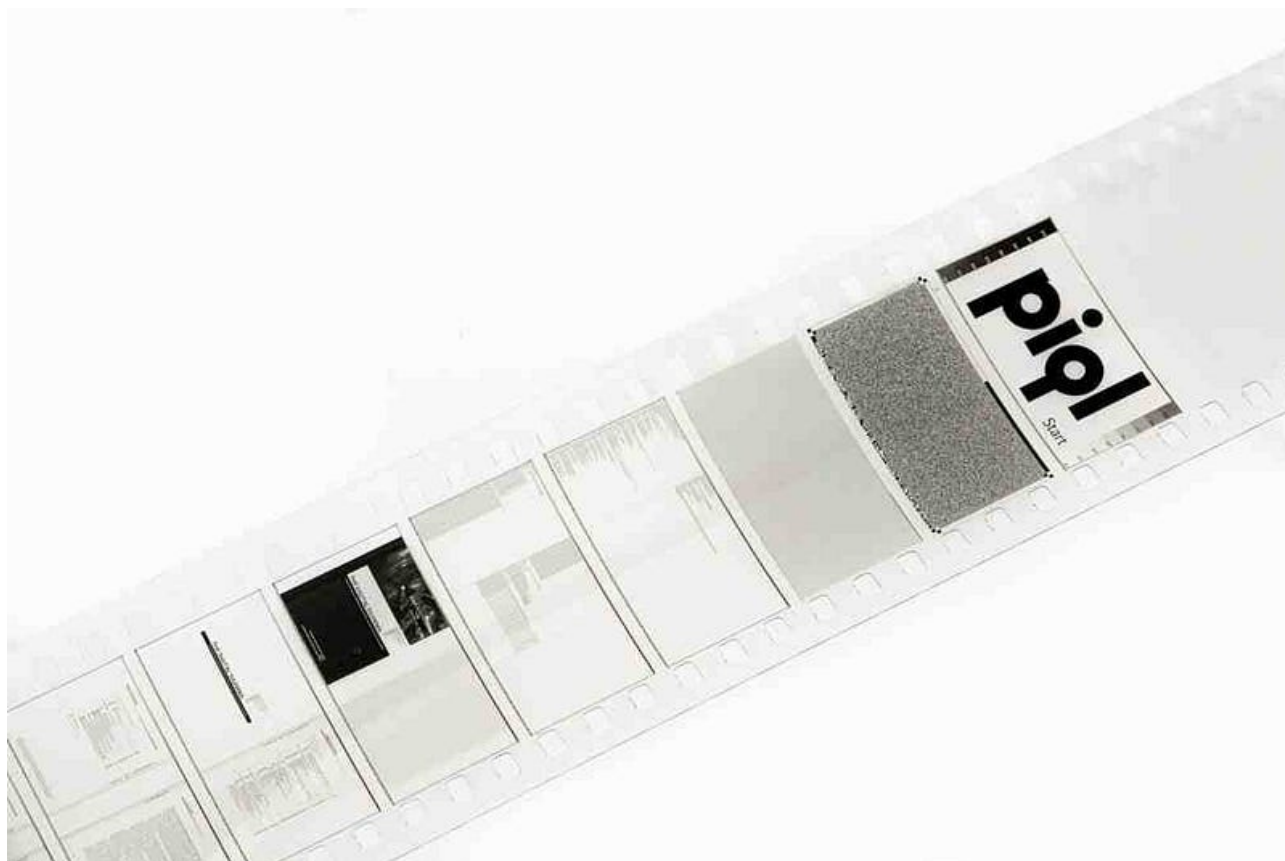
Figure 2:

Dati digitali “analogici” come immagini e microfilm possono essere inseriti normalmente. I dati digitali puri vengono invece codificati in fotogrammi simili a dei QR code che contengono ciascuno un blocco di dati.

Il fatto che la codifica sia “visiva” rende possibile eseguire la decodifica, noto il metodo di codifica, anche senza le apparecchiature originali, usando un oggetto che esegua scansioni ad alta risoluzione ed un computer, dotato di un opportuno software, che riassembli le scansioni nei file digitali originali.

Figure 3:

Alla fine circa un chilometro di pellicola viene inserito in un contenitore appositamente progettato per una lunga conservazione,

Figure 4:

Il periodo di conservazione viene ulteriormente esteso diminuendo la temperatura di conservazione …

… ma per oggi siamo già andati un po' *lunghi*, e quindi qui ci aggiorniamo alla prossima puntata di Archivismi.

Scrivere a Cassandra—Twitter—Mastodon
Videorubrica "Quattro chiacchiere con Cassandra" tempo
Lo Slog (Static Blog) di Cassandra
L'archivio di Cassandra: scuola, formazione e pensiero